

**ANALISIS PERBANDINGAN METODE C4.5, GINI INDEX, DAN NAÏVE BAYES
PADA PEMASARAN KARUNG PT. SUMBER BENGAWAN PLASINDO**



**Disusun sebagai salah satu syarat menyelesaikan Program Studi Strata I pada Jurusan
Informatika Fakultas Komunikasi dan Informatika**

Oleh :

SUCI SURYANI

L200140054

**PROGRAM STUDI INFORMATIKA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
UNIVERSITAS MUHAMMADIYAH SURAKARTA
2018**

HALAMAN PERSETUJUAN

**ANALISIS PERBANDINGAN METODE C4.5, GINI INDEX, DAN NAÏVE
BAYES PADA PEMASARAN KARUNG PT. SUMBER BENGAWAN
PLASINDO**

PUBLIKASI ILMIAH

oleh:

SUCI SURYANI

L 200 140 054

Telah diperiksa dan disetujui untuk diuji oleh:

Dosen Pembimbing



Umi Fadlilah, S.T, M.Eng.

NIP. 197803222005012002

HALAMAN PENGESAHAN

ANALISIS PERBANDINGAN METODE C4.5, GINI INDEX, DAN NAÏVE
BAYES PADA PEMASARAN KARUNG PT. SUMBER BENGAWAN
PLASINDO

OLEH

SUCI SURYANI

L 200 140 054

Telah dipertahankan di depan Dewan Penguji
Fakultas ~~Komunikasi dan Informatika~~
Universitas Muhammadiyah Surakarta
Pada hari ~~Jumat~~ 19 Jan. 2018
dan dinyatakan telah memenuhi syarat

Dewan Penguji:

1. Umi Fadlilah, S.T., M.Eng.
(Ketua Dewan Penguji)
2. Dr. Heru Supriyono, M.Sc.
(Anggota I Dewan Penguji)
3. Aris Rakhmadi, S.T., M.Eng.
(Anggota II Dewan Penguji)

(.....)
(.....)
(.....)

Publikasi ilmiah ini telah diterima sebagai salah satu persyaratan

Untuk memperoleh gelar sarjana

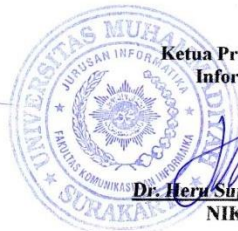
Tanggal 19 Januari 2018.

Mengetahui,



Dekan
Fakultas Komunikasi dan Informatika

Nurgiyatna, M.Sc., Ph.D.
NIK : 881



Ketua Program Studi
Informatika

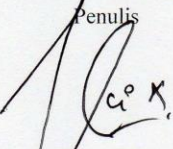
Dr. Heru Supriyono, M.Sc.
NIK : 970

PERNYATAAN

Dengan ini saya menyatakan bahwa dalam naskah publikasi ini tidak terdapat karya yang diajukan untuk memperoleh gelar kesarjanaan di suatu perguruan tinggi dan sepanjang pengetahuan saya juga tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan orang lain, kecuali secara tertulis diacu dalam naskah dan disebutkan dalam daftar pustaka.

Apabila kelak terbukti ada ketidak benaran dalam pernyataan saya diatas, maka akan saya pertanggung jawabkan sepenuhnya.

Surakarta, 16 Januari 2018

Penulis

SUCI SURYANI
L 200 140 054



**UNIVERSITAS MUHAMMADIYAH SURAKARTA
FAKULTAS KOMUNIKASI DAN INFORMATIKA
PROGRAM STUDI INFORMATIKA**

Jl. A Yani Tromol Pos 1 Pabelan Kartasura Telp. (0271)717417, 719483 Fax (0271) 714448
Surakarta 57102 Indonesia. Web: <http://informatika.ums.ac.id>. Email: informatika@ums.ac.id

SURAT KETERANGAN LULUS PLAGIASI

039/A-3-11-3/INF-FKI/1/2018

Assalamu'alaikum Wr. Wb

Biro Skripsi Program Studi Informatika menerangkan bahwa :

Nama : Suci Suryani
NIM : L200140054
Judul : ANALISIS PERBANDINGAN METODE C4.5, GINI INDEX, DAN
NAIVE BAYES PADA PEMASARAN KARUNG PT. SUMBER
BENGAWAN PLASINDO
Program Studi : Informatika
Status : **Lulus**

Adalah benar-benar sudah lulus pengecekan plagiasi dari Naskah Publikasi Skripsi, dengan menggunakan aplikasi Turnitin.

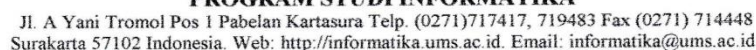
Demikian surat keterangan ini dibuat agar dipergunakan sebagaimana mestinya.

Wassalamu'alaikum Wr. Wb

Surakarta, 30 Januari 2018

Biro Skripsi Informatika

Ihsan Cahyo Utomo, S.Kom., M.Kom.



V

ANALISIS PERBANDINGAN METODE C4.5, GINI INDEX, DAN NAÏVE BAYES PADA PEMASARAN KARUNG PT. SUMBER BENGAWAN PLASINDO

Abstrak

PT. Sumber Bengawan Plasindo atau juga dikenal dengan PT SBP merupakan perusahaan yang bergerak dalam bidang produksi dan distribusi karung plastik yang berlokasi di Kabupaten Karanganyar, Provinsi Jawa Tengah. Karung produksi PT SBP tidak hanya dipasarkan dalam negeri saja, namun telah mencapai pasar luar negeri terbentang dari Asia Tenggara, Eropa hingga Amerika Selatan. Oleh karena itu, tentunya PT SBP memiliki banyak data transaksional, hal ini sangat bermanfaat apabila data tersebut diolah. Jadi peneliti ingin membantu perusahaan dalam pengambilan keputusan untuk memaksimalkan kinerja perusahaan dalam memasarkan produk baik yang diminati maupun yang tidak diminati oleh pelanggan menggunakan metode *C4.5*, *gini index*, dan *naïve bayes*. Variabel yang digunakan dalam proses data mining terdiri dari variabel pemasaran, harga, LPK display name, dan PCS/bale. Dengan proses data mining jumlah produksi harian akan digunakan sebagai sumber strategis untuk menentukan strategi pemasaran perusahaan. Dilihat dari nilai *accuracy*, *precision*, dan *recall* metode *C4.5* dan *gini index* memiliki hasil nilai yang sama dan lebih tinggi dari metode *naïve bayes* yaitu nilai *accuracy* sebesar 96,36%, nilai *precision* sebesar 97,66%, dan nilai *recall* sebesar 92,59%. Sehingga dapat disimpulkan bahwa metode *C4.5* dan *gini index* merupakan metode yang lebih baik dalam penelitian ini. Variabel yang mempengaruhi dalam penelitian ini yaitu variabel harga, kemudian hasil dari metode *C4.5*, *gini index* dan *naïve bayes* produk yang paling banyak diminati oleh pelanggan yaitu produk karung motif.

Kata kunci: *C4.5*, *gini index*, *naïve bayes*, data mining.

Abstract

PT. Sumber Bengawan Plasindo or also known as *PT SBP* is a company which engaged with the production and distribution of plastic sack located in *Karanganyar* district, Central Java. The sack production of *PT SBP* is not only distributed domestically, but have reached the overseas markets stretching from Southeast Asia, Europe to South America. Therefore, *PT SBP* definitely has a lot of transactional data, which is very beneficial if the data are analysed. Thus, the researcher wants to help the company in making the decision to maximize the performance of the company to distribute the products both being interested and un-interested by the costumers by using *C4.5* method, Gini Index, and Naive Bayes. The attributes that are used in data mining process consist of marketing attribute, price, LPK display name and PCS/bale. Through data mining process, the daily production amount will be used as strategic source to decide the marketing strategy of the company. Seen from the accuracy, precision and recall method *C4.5*

and Gini Index have same percentage and been higher than *naïve bayes* method that the percentage of the accuracy is 96, 36%, while its *precision* is 97, 66%, and recall is 92, 59%. Thus can be concluded that C4.5 method and Gini Index is classified as a better method to this study. The variable which influencesto this study is price variable, then the result from C4.5 method, Gini Index, and naïve bayes in which the demanded product by the costumers is the product of motif sack

Key Words: *C4.5, gini index, naïve bayes, mining data.*

1. PENDAHULUAN

PT Sumber Bengawan Plasindo atau PT SBP merupakan perusahaan yang bergerak pada bidang produksi karung yang terletak di Kabupaten Karanganyar, Provinsi Jawa Tengah. PT SBP resmi dibuka pada tahun 2001, hingga pada tahun 2012 berhasil mengekspor produk karung secara masal dalam bentuk lembaran maupun berupa rol ke pasar Asia Tenggara, Eropa hingga Amerika Selatan. Hal ini tentunya tidak luput dari persaingan industri penjualan yang menuntut perusahaan agar memaksimalkan kinerja dalam memasarkan produk, baik yang diminati maupun yang tidak diminati oleh pasar. Minat pasar terhadap produk perusahaan sangat mempengaruhi omset yang akan diperoleh selanjutnya. Dengan demikian perusahaan harus memiliki strategi pemasaran yang lebih tepat dan terarah.

Sementara itu, menurut Husnul (2016) dilihat dari banyaknya jumlah produksi tentunya berdampak pada menumpuknya data, dengan adanya data mining dapat membantu perusahaan dalam mengolah data sehingga dapat memberikan informasi strategis untuk strategi pemasaran produk. Ibnu (2015) mengatakan bahwa strategi pemasaran yang tepat dapat memberikan kepuasan pada konsumen berupa loyalitas, sehingga konsumen menggunakan kembali produk dari perusahaan secara terus menerus.

Meninjau permasalahan di atas, mendorong peneliti untuk membantu perusahaan dalam strategi pemasaran produk dengan membandingkan hasil dari tiga metode data *mining* yaitu C45, gini index, dan *naïve bayes* dengan melakukan penelitian yang berjudul “Analisis Perbandingan Metode C4.5, Gini Index, dan *Naïve Bayes* pada Pemasaran Karung PT Sumber Bengawan Plasindo”. Melalui laporan data produksi karung per hari diharapkan dapat membantu kinerja perusahaan dalam menentukan strategi pemasaran.

2. METODE

2.1 Pengumpulan Data

Penelitian ini menggunakan data produksi harian dari PT Sumber Bengawan Plasindo yang berfokus pada tanggal 10, 11, dan 12 Oktober 2017, dengan jumlah data awal 1071.

1066	TGL PRODUKSI	SHIFT	LPK NO	LPK ITEM NAME	LPK DISPLAY NAME	PCS/BALE	WEIGHT	PEMASARAN
1067	12.10.2017	C3	120	45S124	TRANSPARAN	1000	RINGAN	YA
1068	12.10.2017	C3	121	45S124	TRANSPARAN	1000	RINGAN	YA
1069	12.10.2017	C3	122	75S61-D	MOTIF	500	SEDANG	TIDAK
1070	12.10.2017	C3	123	75S61-D	MOTIF	500	SEDANG	TIDAK
1071	12.10.2017	C3	124	75S61-D	MOTIF	500	SEDANG	TIDAK
1072	12.10.2017	C3	125	75S61-D	MOTIF	500	SEDANG	TIDAK

Gambar 1. Data Awal

Kemudian peneliti melakukan olah data untuk memudahkan dalam perhitungan data mining, yaitu $\frac{2}{3}$ data sebanyak 716 sebagai data training dan $\frac{1}{3}$ data sebanyak 357 sebagai data testing.

2.2 Penentuan Atribut

Metode yang digunakan untuk proses data mining adalah metode C4.5, gini index, dan *naïve bayes* dengan menggunakan atribut sebagai berikut :

Tabel 1. Atribut yang digunakan

Atribut	Variabel
Y	Pemasaran
X1	Harga
X2	LPK Display Name
X3	PCS/Bale
X4	Weight

2.3 Implementasi Data Mining Decision Tree Algoritma C4.5

Rumus *Entropy* yang digunakan sebagai Persamaan 1.

$$Entropy(S) = \sum_b^a - P_b \log_2 P_b \quad (1)$$

Keterangan :

a : nilai pada atribut (kelas)

P_b : jumlah sampel kelas b

Rumus *Information Gain* yang digunakan sebagai Persamaan 2.

$$Gain(S,A) = Entropy(S) - \sum_{ve \text{ Values } (A)} \frac{|Y_v|}{|Y|} Entropy(S_v) \quad (2)$$

2.4 Implementasi Data Mining Gini Index

Rumus Gini Index yang digunakan sebagai Persamaan 3.

$$Gini(A) = 1 - \sum_{k=1}^a P_k^2 \quad (3)$$

Keterangan :

k : Kelas atribut

a : Jumlah kelas pada variable Y

Pk : Jumlah kelas pada atribut k terhadap jumlah kelas a dalam atribut

Rumus Gini Index yang digunakan sebagai persamaan 4.

$$Gini_{split}(A) = \frac{N_1}{N} \cdot Gini(A_1) + \frac{N_2}{N} \cdot Gini(A_2) + \dots + Gini(A_k) \quad (4)$$

2.5 Implementasi Data Mining Naïve Bayes

Metode *naïve bayes* dari teorema sebagai Persamaan 5.

$$P(C|X) = \frac{P(X|C) \cdot P(C)}{P(X)} \quad (5)$$

Keterangan:

X : Data *class* belum diketahui

C : Hipotesis data *class* spesifik

P(C|X) : Probabilitas hipotesis C berdasar kondisi X

P(C) : Probabilitas hipotesis C

P(X|C) : Probabilitas X berdasarkan kondisi hipotesis C

P(X) : Probabilitas X

3. HASIL DAN PEMBAHASAN

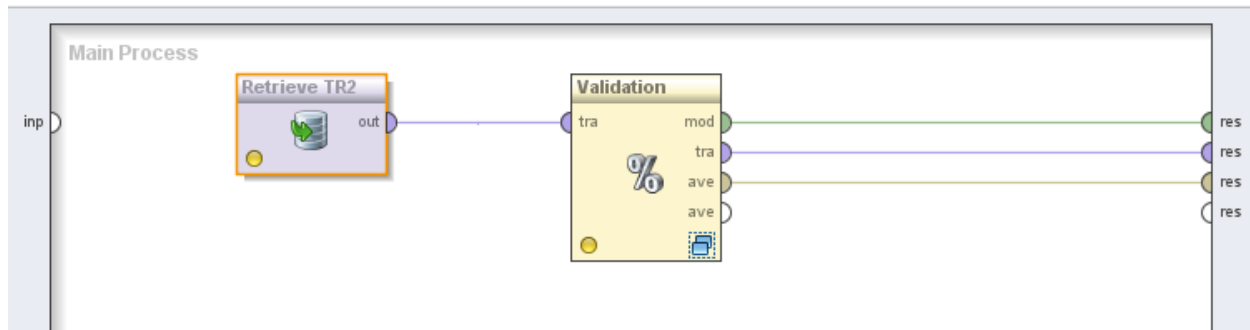
Penelitian ini akan membandingkan hasil dari perhitungan data *mining* yaitu C4.5, gini index, dan *naïve bayes* untuk membantu kinerja perusahaan dalam menentukan strategi pemasaran. Dalam pengolahan data, peneliti menghapus 20 data untuk memudahkan dalam pengkategorian data kemudian peneliti mengkategorikan data menjadi tipe data *polynomial* (3 atau lebih tipe data) dan tipe data *binominal* (2 tipe data) dan telah mendapatkan persetujuan dari PT Sumber Bengawan Plasindo.

HARGA	LPK DISPLAY NAME	PCS/BALE	WEIGHT	PEMASARAN
HARGA ≤ 1500	MOTIF	BANYAK	SEDANG	YA
HARGA ≤ 1500	MOTIF	BANYAK	SEDANG	YA
HARGA ≤ 1500	TRANSPARAN	BANYAK	SEDANG	YA
HARGA ≤ 1500	POLOS	BANYAK	SEDANG	YA
HARGA ≤ 1500	POLOS	BANYAK	SEDANG	YA
1500 > HARGA < 2000	MOTIF	SEDIKIT	SEDANG	YA
1500 > HARGA < 2000	MOTIF	SEDIKIT	SEDANG	YA
HARGA ≤ 1500	MOTIF	BANYAK	SEDANG	YA
1500 > HARGA < 2000	MOTIF	SEDIKIT	SEDANG	YA
1500 > HARGA < 2000	MOTIF	SEDIKIT	SEDANG	YA
HARGA ≥ 2000	POLOS	SEDIKIT	SEDANG	TIDAK
HARGA ≥ 2000	POLOS	SEDIKIT	SEDANG	TIDAK

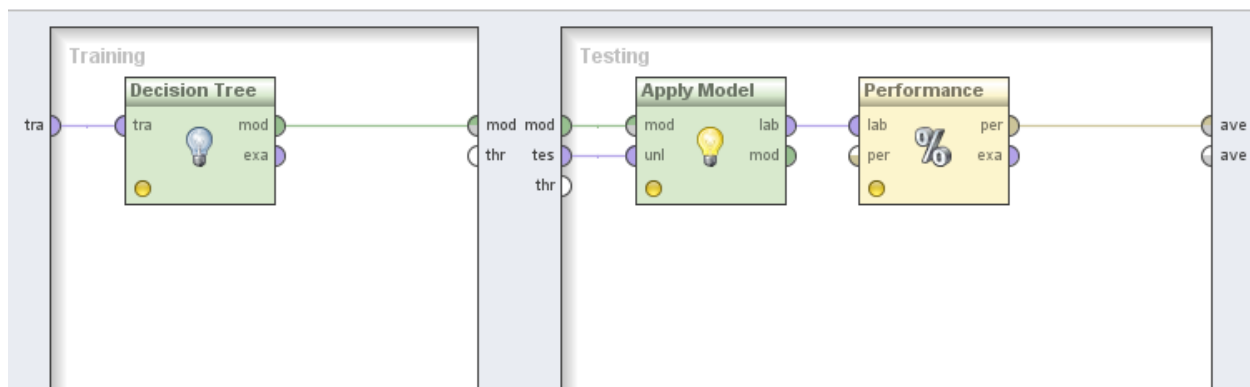
LPK DISPLAY NAME		HARGA
46517	POLOS	56544 TR
3555		40571 TR
60596-B5		455129 TR
75561-B5		565102 TR
56544 TR		60596 JAGO
40571 TR	TRANSPARAN	46517
455129 TR		3555
565102 TR		60596-B5
60596 JAGO		60596 PACK AYAM EMAS
455129 TR		605156 BELAH BALL 500 MTR
565102 TR	MOTIF	50513
60596 JAGO		60533-C GF L-19 SNI
755162 TH SBY 120 JAGO/SL		65538-A
75561 TH BP JAGO		7058
75561 BP JAGO		80514-A
60596 PACK AYAM EMAS	BERWARNA	755162 TH SBY 120 JAGO/SL
755162 TH SBY 120 JAGO/SL		75561 TH BP JAGO
75561 TH BP JAGO		75561 BP JAGO
60596 PACK AYAM EMAS		75561-B5
755162 TH SBY 120 JAGO/SL		75561 TH BP JAGO SL
60533-C GF L-19 SNI	WEIGHT	755162 TH SBY 120 JAGO/SL
65538-A		20-39.9 kg
7058		40-59.9 kg
80514-A		60-80 kg
755162 TH SBY 120 JAGO/SL		
75561 TH BP JAGO	PCS/BALE	500
75561 BP JAGO		1000
75561-B5		SEDIKIT
75561 TH BP JAGO SL		BANYAK
755162 TH SBY 120 JAGO/SL		

Gambar 2. Data yang Telah Diolah

Rancangan proses skema pohon keputusan (*decision tree*) pada metode Algoritma C4.5 dan Gini Index menggunakan Rapid Miner dapat dilihat pada Gambar 3 dan 4.

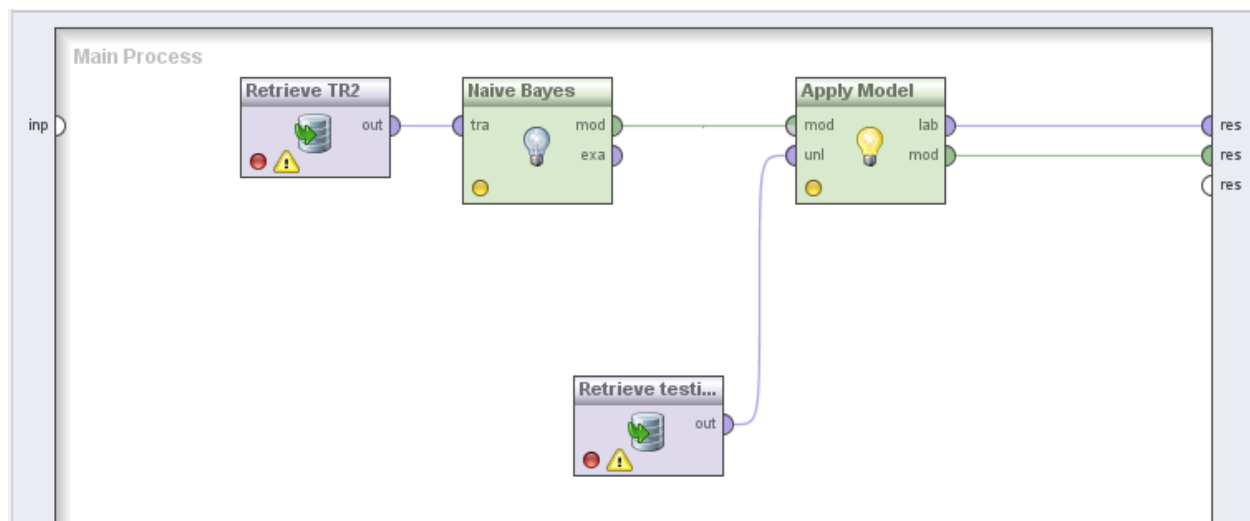


Gambar 3. Rancangan Proses Klasifikasi pada Rapid Miner

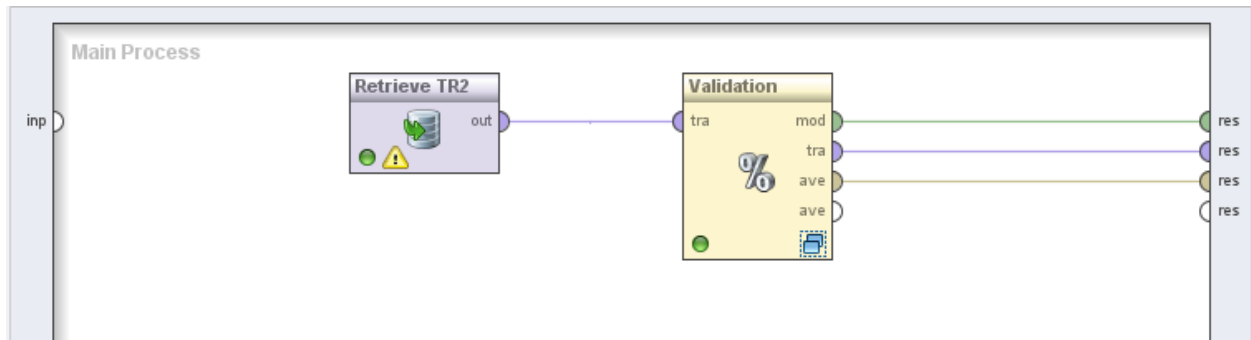


Gambar 4. Rancangan Proses Klasifikasi Decision Tree pada Rapid Miner

Rancangan proses klasifikasi desain metode *naïve bayes* menggunakan Rapid Miner dapat dilihat pada Gambar 5 dan 6.



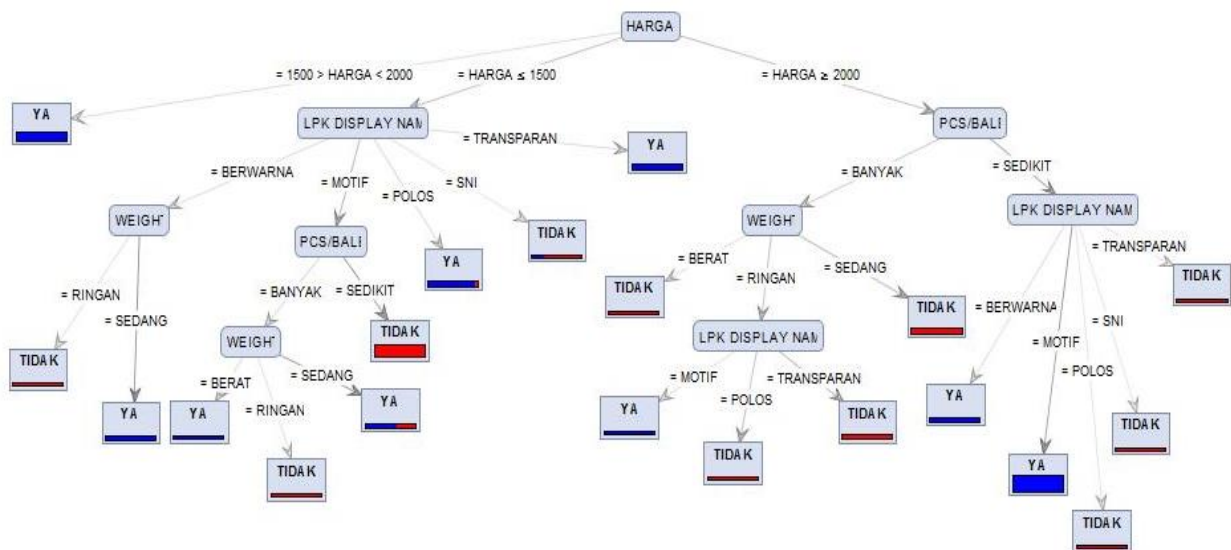
Gambar 5. Desain 1 Metode *Naïve Bayes* pada Rapid Miner



Gambar 6. Desain 2 Metode *Naïve Bayes* pada Rapid Miner

a. Implementasi Data Mining Decision Tree Algoritma C4.5 Menggunakan Rapid Miner

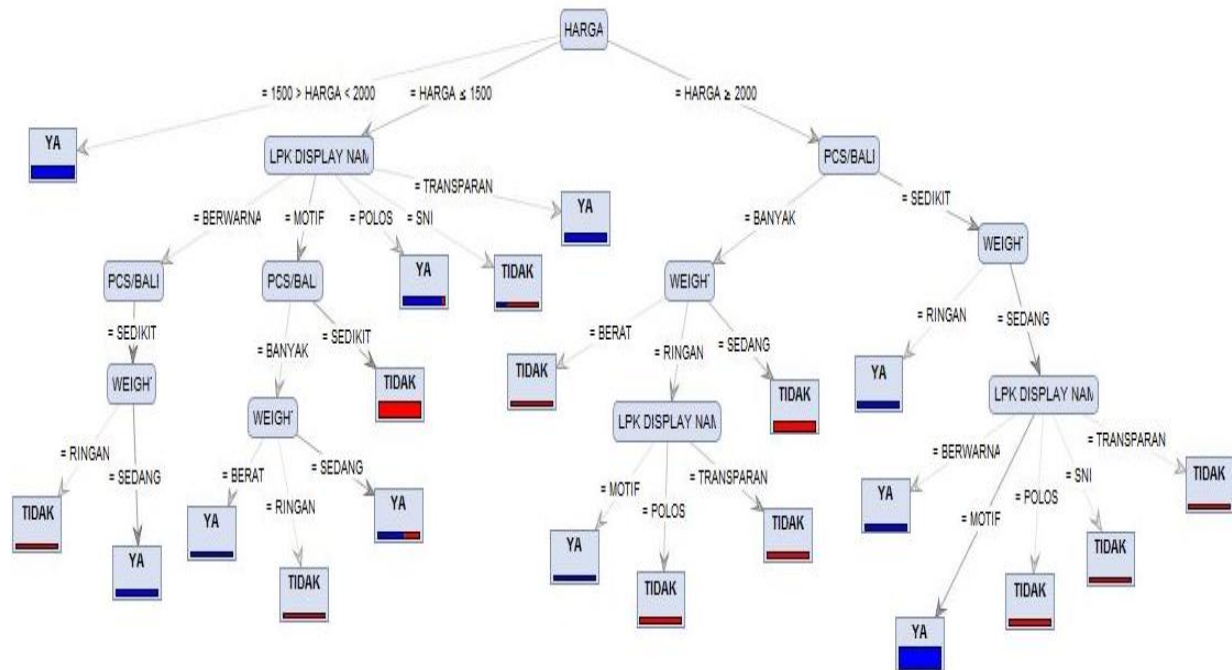
Hasil pohon keputusan dari proses pengolahan data *mining* algoritma C4.5 menggunakan Rapid Miner dapat dilihat pada Gambar 7.



Gambar 7. Hasil Pohon Keputusan Algoritma C4.5

b. Implementasi Data Mining Decision Tree Gini Index Menggunakan Rapid Miner

Hasil pohon keputusan dari proses pengolahan data *mining* algoritma Gini Index menggunakan Rapid Miner dapat dilihat pada Gambar 8.



Gambar 8. Hasil Pohon Keputusan Algoritma Gini Index

Terdapat perbedaan hasil pohon keputusan algoritma C4.5 dan gini index yaitu pada harga ≤ 1500 simpul cabang (internal node) LPK display name algoritma gini index terdapat penambahan simpul cabang PCS/Bale. Serta pada harga ≥ 2000 simpul cabang PCS/Bale terdapat penambahan simpul cabang weight.

c. Implementasi Data Mining Algoritma Naïve Bayes Menggunakan Rapid Miner 5

Hasil prediksi data testing dan tingkat *confidence* pada masing-masing proses perhitungan data metode *naïve bayes* menggunakan Rapid Miner dapat dilihat pada Gambar 9.

Row No.	confidence(YA)	confidence(TIDAK)	prediction(PEMASARAN)	HARGA	LPK DISPLAY NAME	PCS/BALE	WEIGHT
1	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
2	0.519	0.481	YA	HARGA \leq 1500	MOTIF	SEDIKIT	SEDANG
3	0.397	0.603	TIDAK	HARGA \leq 1500	POLOS	SEDIKIT	SEDANG
4	0.599	0.401	YA	HARGA \geq 2000	POLOS	SEDIKIT	SEDANG
5	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
6	1.000	0.000	YA	1500 > HARGA < 2000	MOTIF	SEDIKIT	SEDANG
7	1.000	0.000	YA	1500 > HARGA < 2000	TRANSPARAN	BANYAK	RINGAN
8	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
9	0.346	0.654	TIDAK	HARGA \geq 2000	POLOS	BANYAK	BERAT
10	0.519	0.481	YA	HARGA \leq 1500	MOTIF	SEDIKIT	SEDANG
11	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
12	0.519	0.481	YA	HARGA \leq 1500	MOTIF	SEDIKIT	SEDANG
13	0.498	0.502	TIDAK	HARGA \geq 2000	TRANSPARAN	BANYAK	RINGAN
14	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
15	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
16	0.519	0.481	YA	HARGA \leq 1500	MOTIF	SEDIKIT	SEDANG
17	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
18	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
19	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
20	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
21	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
22	0.710	0.290	YA	HARGA \geq 2000	MOTIF	SEDIKIT	SEDANG
23	0.538	0.462	YA	HARGA \geq 2000	TRANSPARAN	BANYAK	SEDANG
24	0.538	0.462	YA	HARGA \geq 2000	TRANSPARAN	BANYAK	SEDANG

Gambar 9. Hasil ExampleSet *Naïve Bayes*

d. Implementasi Algoritma C4.5

Membandingkan hasil perhitungan X1, X2, X3, X4 dimana nilai tertinggi digunakan sebagai simpul akar (*root node*).

Tabel 2. Hasil Perhitungan C45

Atribut	Variabel	Nilai C4.5
X1	Harga	0,115
X2	LPK Display Name	0,049
X3	PCS/Bale	0,015
X4	Weight	0,001

Dilihat dari tabel 2, dimana Harga merupakan nilai tertinggi, maka dijadikan sebagai simpul akar (*root node*).

e. Implementasi Gini Index

Membandingkan hasil perhitungan X1, X2, X3, X4 dimana nilai terendah digunakan sebagai simpul akar (*root node*).

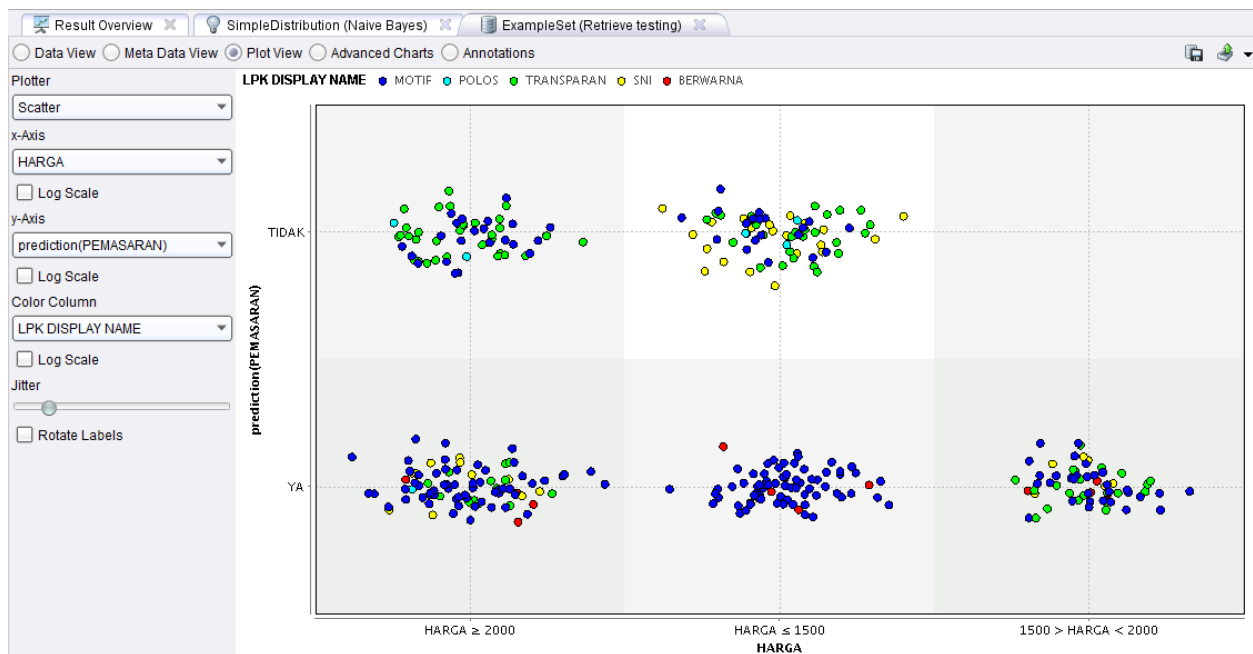
Tabel 3. Perhitungan Gini Index

Atribut	Variabel	Nilai Gini Index
X1	Harga	0,414
X2	LPK Display Name	0,444
X3	PCS/Bale	0,460
X4	Weight	0,469

Dilihat dari tabel 2, dimana Harga merupakan nilai terendah, maka dijadikan sebagai simpul akar (*root node*).

3.6 Implementasi *Naïve Bayes*

Hasil isi data training metode *naïve bayes* dilihat dari plot view pada ExampleSet dapat dilihat pada Gambar 10.



Gambar 10. Hasil Plot View *Naïve Bayes* pada Rapid Miner

Hasil dari plot view *naïve bayes* pada Gambar 10 dapat disimpulkan bahwa produk karung motif pada harga ≥ 2000 , harga ≤ 1500 , dan $1500 > \text{harga} < 2000$ merupakan produk karung yang sering diproduksi dan dipesan oleh pelanggan. Sedangkan produk karung polos pada harga ≥ 2000 merupakan produk yang kurang diminati oleh pelanggan, hal ini dilihat dari jumlah produk yang diproduksi. Pada harga ≤ 1500 produk karung motif, polos, dan SNI juga merupakan produk yang kurang diminati oleh pelanggan.

3.7 Perbandingan 3 Metode

Melakukan perbandingan dari hasil analisis perhitungan ketiga metode data *mining* yaitu C45, Gini Index, dan *Naïve Bayes* dengan melihat hasil nilai *accuracy*, *precision*, dan *recall*.

Secara umum *accuracy*, *precision*, dan *recall* dapat dirumuskan sebagai berikut:

		Nilai Asli	
		TRUE	FALSE
Nilai Prediksi	TRUE	TP (True Positive) Corect result	FP (False Positive) Unexpect result
	FALSE	FN (False Negative) Missing result	TN (True Negative) Corect absence of result

a. $accuracy = \frac{TP}{TP+TN+FP+FN}$

b. $precision = \frac{TP}{TP+FP}$

c. $recal = \frac{TP}{TP+FN}$

Tabel 4. Perbandingan 3 Metode

Metode	<i>Accuracy</i>	<i>Precision</i>	<i>Recall</i>
C4.5	96,36%	97,66%	92,59%
Gini Index	96,36%	97,66%	92,59%
Naïve Bayes	57,20%	40,11%	27,04%

Hasil perbandingan 3 metode pada Tabel 4 dapat disimpulkan bahwa metode C4.5 dan Gini Index memiliki nilai *accuracy*, *precision*, dan *recall* yang sama dan lebih tinggi sehingga metode tersebut merupakan metode yang lebih baik dalam penelitian ini. Nilai *accuracy*, *precision*, dan *recall* pada metode *naïve bayes* lebih rendah, maka dapat disimpulkan ketepatan metode memiliki nilai keakuratan rendah.

3.8 Intrepretasi Hasil Penelitian

Hasil dari analisis penelitian menggunakan metode C4.5 dan Gini Index menghasilkan algoritma *decision tree* sedangkan pada metode *naïve bayes* menghasilkan prediksi terhadap tingkat *confidence* pada masing-masing perhitungan data.

3.8.1 Intepretasi Algoritma C4.5

Berdasarkan Gambar.7 dalam analisis penelitian ini atribut harga memiliki nilai tertinggi sehingga harga merupakan variable yang paling mempengaruhi, terbukti atribut harga merupakan simpul akar (*root node*). Aturan yang terbentuk dari hasil penelitian metode C4.5 adalah:

1. Seseorang akan berminat membeli produk (YA) jika kondisi sebagai berikut:
 - a) $\text{Harga} = 1500 > \text{Harga} < 2000$, (kondisi lain diabaikan).
 - b) $\text{Harga} = \leq 1500$, LPK Display Name = Berwarna, Weight = Sedang.
 - c) $\text{Harga} = \leq 1500$, LPK Display Name = Motif, PCS/Bale = Banyak, Weight = Berat.
 - d) $\text{Harga} = \leq 1500$, LPK Display Name = Motif, PCS/Bale = Banyak, Weight = Sedang.
 - e) $\text{Harga} = \leq 1500$, LPK Display Name = Polos, (nilai atribut lain diabaikan).
 - f) $\text{Harga} = \leq 1500$, LPK Display Name = Transparan, (nilai atribut lain diabaikan).
 - g) $\text{Harga} = \geq 2000$, PCS/Bale = Banyak, Weight = Ringan, LPK Display Name = Motif.
 - h) $\text{Harga} = \geq 2000$, PCS/Bale = Sedikit, LPK Display Name = Berwarna.
 - i) $\text{Harga} = \geq 2000$, PCS/Bale = Sedikit, LPK Display Name = Motif.
2. Seseorang kurang berminat membeli produk (TIDAK) jika kondisi sebagai berikut:
 - a) $\text{Harga} = \leq 1500$, LPK Display Name = Berwarna, Weight = Ringan.
 - b) $\text{Harga} = \leq 1500$, LPK Display Name = Motif, PCS/Bale = Banyak, Weight = Ringan.
 - c) $\text{Harga} = \leq 1500$, LPK Display Name = Motif, PCS/Bale = Sedikit.
 - d) $\text{Harga} = \leq 1500$, LPK Display Name = SNI, (nilai atribut lain diabaikan).
 - e) $\text{Harga} = \geq 2000$, PCS/Bale = Banyak, Weight = Berat, (nilai atribut lain diabaikan).
 - f) $\text{Harga} = \geq 2000$, PCS/Bale = Banyak, Weight = Ringan, LPK Display Name = Polos.
 - g) $\text{Harga} = \geq 2000$, PCS/Bale = Banyak, Weight = Ringan, LPK Display Name = Transparan.
 - h) $\text{Harga} = \geq 2000$, PCS/Bale = Banyak, Weight = Sedang, (nilai atribut lain diabaikan).
 - i) $\text{Harga} = \geq 2000$, PCS/Bale = Sedikit, LPK Display Name = Polos, (nilai atribut lain diabaikan).

- j) Harga = ≥ 2000 , PCS/Bale = Sedikit, LPK Display Name = SNI, (nilai atribut lain diabaikan).
- k) Harga = ≥ 2000 , PCS/Bale = Sedikit, LPK Display Name = Transparan, (nilai atribut lain diabaikan).

3.8.2 Intepretasi Algoritma Gini Index

Berdasarkan Gambar.8 dalam analisis penelitian ini atribut harga memiliki nilai terendah sehingga harga merupakan variable yang paling mempengaruhi, terbukti atribut harga merupakan simpul akar (*root node*). Aturan yang terbentuk dari hasil penelitian metode Gini Index adalah:

1. Seseorang akan berminat membeli produk (YA) jika kondisi sebagai berikut:
 - a) Harga = $1500 > \text{Harga} < 2000$, (kondisi lain diabaikan).
 - b) Harga = ≤ 1500 , LPK Display Name = Berwarna, PCS/Bale = Sedikit, Weight = Sedang.
 - c) Harga = ≤ 1500 , LPK Display Name = Motif, PCS/Bale = Banyak, Weight = Berat.
 - d) Harga = ≤ 1500 , LPK Display Name = Motif, PCS/Bale = Banyak, Weight = Sedang.
 - e) Harga = ≤ 1500 , LPK Display Name = Polos, (nilai atribut lain diabaikan)
 - f) Harga = ≤ 1500 , LPK Display Name = Transparan, (nilai atribut lain diabaikan).
 - g) Harga = ≥ 2000 , PCS/Bale = Banyak, Weight = Ringan, LPK Display Name = Motif.
 - h) Harga = ≥ 2000 , PCS/Bale = Sedikit, Weight = Ringan.
 - i) Harga = ≥ 2000 , PCS/Bale = Sedikit, Weight = Sedang, LPK Display Name = Berwarna.
 - j) Harga = ≥ 2000 , PCS/Bale = Sedikit, Weight = Sedang, LPK Display Name = Motif.
2. Seseorang kurang berminat membeli produk (TIDAK) jika kondisi sebagai berikut:
 - a) Harga = ≤ 1500 , LPK Display Name = Berwarna, PCS/Bale = Sedikit, Weight = Ringan.
 - b) Harga = ≤ 1500 , LPK Display Name = Motif, PCS/Bale = Banyak, Weight = Ringan.
 - c) Harga = ≤ 1500 , LPK Display Name = Motif, PCS/Bale = Sedikit.
 - d) Harga = ≤ 1500 , LPK Display Name = SNI, (nilai atribut lain diabaikan).
 - e) Harga = ≥ 2000 , PCS/Bale = Banyak, Weight = Berat.
 - f) Harga = ≥ 2000 , PCS/Bale = Banyak, Weight = Ringan, LPK Display Name = Polos.
 - g) Harga = ≥ 2000 , PCS/Bale = Banyak, Weight = Ringan, LPK Display Name = Transparan.
 - h) Harga = ≥ 2000 , PCS/Bale = Banyak, Weight = Sedang, (nilai atribut lain diabaikan).
 - i) Harga = ≥ 2000 , PCS/Bale = Sedikit, Weight = Sedang, LPK Display Name = Polos.

- j) Harga = ≥ 2000 , PCS/Bale = Sedikit, Weight = Sedang, LPK Display Name = SNI.
- k) Harga = ≥ 2000 , PCS/Bale = Sedikit, Weight = Sedang, LPK Display Name = Transparan.

3.8.2 Interpretasi Algoritma *Naïve Bayes*

Berdasarkan Gambar. 9 dalam analisis penelitian ini mengambil kesimpulan dari metode *naïve bayes* yaitu :

1. Harga ≥ 2000
 - a) YA = LPK Display Name Motif merupakan produk yang banyak diminati.
 - b) TIDAK = LPK Display Name Transparan merupakan produk yang tidak/kurang diminati.
2. Harga ≤ 1500
 - a) YA = LPK Display Name Motif merupakan produk yang banyak diminati.
 - b) TIDAK = LPK Display Name SNI merupakan produk yang tidak/kurang diminati.
3. $1500 > \text{Harga} < 2000$
 - a) YA = Hampir semua LPK Display Name sama banyak diminati.

4. PENUTUP

4.1 Kesimpulan

Berdasarkan hasil analisis perbandingan 3 metode yaitu C4.5, Gini Index, dan *Naïve Bayes* dapat ditarik kesimpulan sebagai berikut :

1. Berdasarkan hasil perbandingan metode C4.5 dan Gini Index variabel yang paling berpengaruh terhadap penjualan adalah variabel harga. Serta memiliki nilai *accuracy* 96,36%, nilai *precision* 97,66%, dan nilai *recall* 92,59% yang sama.
2. Berdasarkan hasil perhitungan metode *naïve bayes* menggunakan rapid miner nilai *accuracy* 57,20%, nilai *precision* 40,11%, dan nilai *recall* 27,04% merupakan nilai terendah dari kedua metode yang lain yaitu metode C4.5 dan Gini Index.
3. Berdasarkan hasil perbandingan metode C4.5, Gini Index, dan *Naïve Bayes*. Metode C4.5 dan Gini Index memiliki nilai *accuracy*, *precision*, dan *recall* yang lebih tinggi dibanding nilai *accuracy*, *precision*, dan *recall* *Naïve Bayes* sehingga metode C4.5 dan Gini Index merupakan metode yang lebih baik dalam penelitian ini.
4. Produk yang sering diproduksi dan dipesan oleh pelanggan yaitu produk karung motif.

4.2 SARAN

Adapun saran agar penelitian lebih baik dari penelitian sebelumnya yaitu:

1. Agar melakukan penelitian dengan membandingkan lebih banyak metode dan menggunakan metode yang berbeda dari penelitian sebelumnya.
2. Penelitian ini tidak terimplementasi pada sebuah system aplikasi, maka perlu dibuat sistem aplikasi web.
3. Menggunakan Rapid Miner versi terbaru yang lebih lengkap tampilannya.

DAFTAR PUSTAKA

- Asmara, H. M. (2016). Analisa Perbandingan Hasil Pohon Keputusan Dengan Gain Ratio, Information Gain, Gini Index pada Pemasaran Produk Herbal di CV Al-Ghuroba. Surakarta: Universitas Muhammadiyah Surakarta.
- Rohman, I. F. (2015). Penerapan Algoritma C4.5 pada Kepuasan Pelanggan Perum Damri. Semarang: Universitas Dian Nuswantoro Semarang.
- Santoso, T. B. (2014). Analisa dan Penerapan Metode C4.5 untuk Prediksi Loyalitas Pelanggan. *Jurnal Ilmiah Fakultas Teknik LIMIT'S*, 10.
- Gunawan, R., & Mustofa, K. (2016). Pencarian Aturan Naive Bayes Sematic Web untuk Obat Tradisional Indonesia. *JNTETI*, 5, 192-200. Yogyakarta: Universitas Gajah Mada.
- Nugroho, Y.S. (2016). Modul Praktikum Data Warehouse dan Data Mining. Surakarta: Universitas Muhammadiyah Surakarta.
- Nurjaya, W., & Adani, Y. Penerapan Algoritma Naïve Bayes untuk Memprediksi Keputusan Nasabah Telemarketing Dalam Menawarkan Deposito. Bandung: STMIK LPKIA Bandung.
- Usman, T. (2016). Perbandingan 3 Metode Data Mining untuk Penentuan Hipertensi di Rumah Sakit Umum Daerah Dr. Moewardi Surakarta. Surakarta: Universitas Muhammadiyah Surakara.
- Deepa, S., & Vijayarani, S. (2014). Naive Bayes Classification for Predicting Disease in Haemoglobin Protein Sequences. *International Journal of Computational Intelligence and Informatics*, 3, 278-283.